



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Phylogenetic reclassification of vertebrate melatonin receptors to include Mel1d

Citation for published version:

Denker, E, Ebesson, LOE, Hazlerigg, DG & Macqueen, D 2019, 'Phylogenetic reclassification of vertebrate melatonin receptors to include Mel1d', *G3: Genes | Genomes | Genetics*.
<https://doi.org/10.1534/g3.119.400170>

Digital Object Identifier (DOI):

[10.1534/g3.119.400170](https://doi.org/10.1534/g3.119.400170)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

G3: Genes | Genomes | Genetics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Phylogenetic reclassification of vertebrate melatonin receptors to include Mel1d

Elsa Denker ^{*}, Lars O. E. Ebbesson ^{*}, David G. Hazlerigg [†], Daniel J. Macqueen [‡]

^{*} *NORCE Environment, NORCE Norwegian Research Centre, 5008 Bergen, Norway*

[†] *Department of Arctic and Marine Biology, UiT: the Arctic University of Norway, 9037 Tromsø, Norway*

[‡] *The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, EH25 9RG Edinburgh, United Kingdom*

13

14 Running title: Mel1d - a new vertebrate melatonin receptor

15

16 Keywords: Melatonin receptor; Mel1d; vertebrate evolution; phylogenetic analysis;
17 comparative genomics; conserved synteny; teleost fish

18

19 Corresponding authors:

20 Daniel J. Macqueen, The Roslin Institute, University of Edinburgh, Easter Bush

21 Campus, Midlothian, EH25 9RG, United Kingdom. Phone: (+44) 131 651 9249.

22 Email: daniel.macqueen@roslin.ed.ac.uk

23

24 Elsa Denker, Integrative Fish Biology Group, NORCE Environment, NORCE

25 Norwegian Research Center, Thormøhlensgate 55, 5008 Bergen, Norway. Phone:
26 (+47) 55 58 44 16. Email: elsa.denker@uib.no

27

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

ABSTRACT

The circadian and seasonal actions of melatonin are mediated by high affinity G-protein coupled receptors (melatonin receptors, MTRs), classified into phylogenetically distinct subtypes based on sequence divergence and pharmacological characteristics. Three vertebrate MTR subtypes are currently described: MT1 (MTNR1A), MT2 (MTNR1B), and Mel1c (MTNR1C / GPR50), which exhibit distinct affinities, tissue distributions and signaling properties. We present phylogenetic and comparative genomic analyses supporting a revised classification of the vertebrate MTR family. We demonstrate four ancestral vertebrate MTRs, including a novel molecule hereafter named Mel1d. We reconstructed the evolution of each vertebrate MTR, detailing genetic losses in addition to gains resulting from whole genome duplication events in teleost fishes. We show that Mel1d was lost separately in mammals and birds and has been previously mistaken for an MT1 paralogue. The genetic and functional diversity of vertebrate MTRs is more complex than appreciated, with implications for our understanding of melatonin actions in different taxa. The significance of our findings, including the existence of Mel1d, are discussed in an evolutionary and functional context accommodating a robust phylogenetic assignment of MTR gene family structure.

49

50

INTRODUCTION

51

52 Melatonin is an ancient eukaryotic signalling molecule that regulates diverse
53 biological functions. While best known as a regulator of biological rhythms in
54 humans, this hormone also regulates energy balance, temperature, behavior, blood
55 pressure, and seasonal reproduction. Melatonin is secreted by the pineal gland and
56 targets the brain as well as peripheral tissues (Hardeland *et al.* 2011, Slominski *et al.*
57 2012), but is also produced by several tissues, eliciting paracrine effects (Weaver
58 and Reppert 1990). The actions of melatonin depend on the spatiotemporal
59 expression of high-affinity melatonin receptors (MTR), representing a specific class of
60 G protein-coupled receptor (GPCR).

61

62 Three paralogous MTR family members have been characterized in jawed
63 vertebrates, namely MT1 (Mel1a / MTNR1A), MT2 (Mel1b / MTNR1B), and Mel1c
64 (MTNR1C / GPR50 in mammals) (Reppert *et al.* 1994, 1995a, 1995b). Despite
65 showing overlap in expression, these different MTRs have evolved unique functions.
66 MT1 has a higher affinity for melatonin than MT2 (Dubocovich and Markowska 2005),
67 and in mammals, Mel1c has lost the ability to bind melatonin (Dufourny *et al.* 2008),
68 though it does modulate melatonin signaling via its association with MT1 (Levoye *et al.*
69 2006). While MT1 associates with a range of G proteins to activate several distinct
70 signalling pathways, eliciting wide-ranging cellular effects (Witt-Enderby *et al.* 2003),
71 MT2 associates with a single G protein (Jockers *et al.* 2008). Owing to such
72 functional divergence, different MTRs may have very distinct biological effects, even
73 when expressed in the same cell types (e.g. Dubocovich and Markowska 2005).

74

A past study demonstrated melatonin binding in the brain of jawed vertebrates and lamprey, but not in hagfishes or amphioxus (Vernadakis *et al.* 1998). Thus, it is likely that high-affinity MTRs were present in the vertebrate ancestor, and were secondarily lost in some jawless fishes, as noted for several other traits (e.g. reduction of vertebrae-like elements - Ota *et al.* 2011; *Dlx* genes - Sugahara *et al.* 2013; reviewed in Kuraku 2013). MTR-like GPCR genes have also been discovered in urochordates, cephalochordates, hemichordates and echinoderms (Kamesh *et al.* 2008, Nordstrom *et al.* 2008, Krishnan *et al.* 2013), but their evolutionary affinity to the vertebrate MTRs remains ambiguous. The distinct MTRs of jawed vertebrates potentially originated during two rounds (2R) of whole genome duplication (WGD) at the stem of vertebrate evolution (e.g. Dehal and Boore, 2005), though this is yet to be established. Additional expansions in the MTR family of fishes (e.g. Shang & Zhdanova 2007; Hong *et al.* 2014) may owe to a further round of teleost-specific WGD ('Ts3R') in the common teleost ancestor, or additional lineage-specific WGD events in some lineages, e.g. the salmonid-specific 4R ('Ss4R') (Macqueen and Johnston, 2014; Lien *et al.* 2016), though, again, this has not been properly explored.

The overarching goal of this study was to re-examine the evolutionary history of vertebrate MTRs, using data in publically-available sequence databases for robust phylogenetic and comparative genomic reconstructions. Our findings concretely demonstrate a fourth ancestral MTR ('Mel1d'), along with teleost-specific expansions in MTR diversity, likely owing to Ts3R and Ss4R. With a new evolutionary framework in place we reinterpret findings on vertebrate MTR sequence divergence and expression from past studies. Overall, this study highlights substantial unexplored diversity in MTR signalling within vertebrates, pointing to new lines of investigation.

MATERIALS AND METHODS

Sequence and phylogenetic analyses

Amino acid sequences encoded by MTR or FAT protocadherin family member genes were collected from representative jawed vertebrate species with high-quality genome assemblies. Details of these sequences are given in Table S1 (MTR) and Table S2 (FAT), which include database accession numbers and nomenclature matching the findings of our phylogenetic analyses. As a start point for the analysis, MTR/FAT proteins of human (*i.e.* MT1/MT2/Mel1c/GPR50 or FAT1/2/3) were used as queries in BLASTp (Altschul *et al.* 1997) searches to identify homologues within the NCBI database (<https://www.ncbi.nlm.nih.gov/>). We also used the Ensembl genome browser (<https://www.ensembl.org/>) to collect MTR family proteins from several species, using the EnsemblCompara method (Vilella *et al.* 2009).

The sequences were aligned using MAFFT v.7 (Katoh and Standley, 2013) with default settings and subjected to quality filtering using GBlocks with default settings (Talavera and Castresana, 2007). Final alignments of 300 (MTR) and 2,540 (FAT) amino acid positions (Additional Dataset 1) were used for tree building, done using Bayesian (BY) and maximum likelihood (ML) (MTR) or just ML (FAT) methods. ML trees were generated using IQ-TREE (Nguyen *et al.* 2015) via the IQ-TREE webserver (Trifinopoulos *et al.* 2016), employing the best-fitting amino acid substitution model selected with ModelFinder (Kalyaanamoorthy *et al.* 2017) under the Bayesian information criterion. The best fit models were JTT+F+I+ G4 for MTR and JTT+G4+I for FAT, where 'JTT' is the general matrix of Jones *et al.* 1992, '+I' includes empirical estimation of the proportion of invariant sites, '+F' includes empirical estimation of amino acid frequencies and '+G4' denotes estimation of the

gamma distribution parameter with 4 rate classes. The stability of branching in the ML trees was assessed using 1,000 ultrafast bootstrap replicates, (Hoang *et al.* 2018). The BY analysis (MTR dataset) was done in BEAST v1.8.3 (Drummond *et al.* 2012), employing an uncorrelated relaxed clock model (Drummond *et al.* 2005) and a Yule speciation prior (Gernhard, 2008), along with the best-fitting substitution model selected by IQ-TREE. A Markov chain Monte Carlo (MCMC) chain of 50 million generations was generated and sampled every 5,000 generations. Convergence of the MCMC chain was assessed using Tracer v1.7.1 <http://beast.bio.ed.ac.uk/tracer>. A maximum clade credibility tree was generated in TreeAnnotator (Drummond *et al.* 2012) after removal of the first 10% sampled trees.

Comparative genomic and sequence analyses

Synteny analyses were performed using Ensembl genome browser annotations via the Genomicus platform (Nguyen *et al.* 2018). These analyses were supplemented with data from NCBI GenBank for species not available in Ensembl. Gene prediction and annotation for *Lethenteron camtschaticum* was performed using FGENESH (Solovyev *et al.* 2006). Comparative analyses of MTR family amino acid sequences was done using the final alignment described above (note: the Gblocks filtering step served to remove flanking regions outside the transmembrane/loop regions, which were unaltered). The sequence similarity of the proposed vestigial MTR-like pseudogenes identified in our synteny analyses was established using BLASTx within the NCBI database.

Data Availability

Supplemental material described in the paper is available at Figshare: <https://gsajournals.figshare.com/s/56f29e83cc0ec8748842>. Fig. S1. ML phylogenetic

analysis of MTRs in vertebrates. This analysis was done using IQ-TREE with a high-confidence alignment of eighty MTRs (300 amino acid positions; Additional Dataset 1) and the best-fitting amino acid substitution model (JTT+F+I+G4). Numbers on branches are bootstrap support values. Other details as in the Fig. 1 legend (see main text) Table S1 provides details of all protein sequences used for phylogenetic analyses of the vertebrate MTR family. Table S2 provides details of all sequences used for phylogenetic analyses of the vertebrate FAT protocadherin family. Additional Dataset 1 is the MTR sequence alignment used for phylogenetic analysis and comparative sequence analysis. Additional Dataset 2 is the FAT alignment used for phylogenetic analysis.

RESULTS

Four MTRs are retained in jawed vertebrates

We identified eighty unique MTR family member proteins in sequence databases representing a standardized set of eighteen jawed vertebrate lineages (see MATERIALS AND METHODS). A phylogenetic analysis was done using a BY method (Fig. 1) incorporating a relaxed molecular clock model, which allows estimation of the most plausible root location in the tree (Drummond *et al.* 2006; *e.g.* Macqueen and Wilcox 2014; Redmond *et al.* 2018). Four distinct MTR clades (Fig. 1) had strong statistical support (posterior probability, PP: >0.96), and each was represented by cartilaginous fish, as well as ray-finned and lobe-finned fish lineages, with branching patterns closely matching expected species phylogeny (Fig. 1). Three of these clades correspond to known ancestral vertebrate MTR family members (*e.g.* Dufourny *et al.* 2008). The fourth clade is hereafter named 'Mel1d'. The same four clades were strongly supported in an unrooted ML phylogenetic analysis (bootstrap support: >96%) congruent with the BY tree (Fig. S1).

These analyses indicate that four distinct MTRs existed in the jawed vertebrate ancestor. However, the phylogenetic affinity of the four MTRs remains equivocal in the BY analysis, with moderate support for Mel1d/MT1 (PP: 0.87) and MT2/Mel1C (PP: 0.53) being paralogues, which can be explained parsimoniously by 2R (Fig. 1).

Evolutionary history of individual vertebrate MTRs

Expanding on the above findings, we reconstructed a more detailed evolutionary history for each ancestral MTR in jawed vertebrates, accommodating gene losses, in addition to gains resulting from WGD events in teleosts (summarized in Fig. 2).

Mel1d was encoded by a single gene in all represented species (Fig. 1, Fig. 2a) including teleosts, consistent with the loss of any paralogues created during Ts3R and Ss4R. In lobe-finned fish, Mel1d was identified in a coelacanth, an amphibian, and two reptiles, but was not identified in the mammals and birds represented in our trees (Fig. 2a). As only a small number of bird and mammals were included, we decided to search more broadly for Mel1d orthologues. Hence, BLAST searches of the complete set of proteins stored in NCBI for mammals (~4.6 million) and birds (~2.8 million) were done using reptile Mel1d orthologues as the query. Though hundreds of bird and mammal genomes are available in NCBI with protein-level annotations (spanning the diversity of each lineage), the top mammal/bird hit for reptile Mel1D was always MT1/MTNR1A (not shown). Considering our current understanding of amniote phylogeny (*e.g.* Chiari *et al.* 2012), our data requires that independent losses of Mel1d occurred in the ancestors to birds and mammals.

For all studied vertebrate species outside teleosts, we identified one copy of MT1, barring spotted gar, where MT1 was not identified (Fig. 1, Fig. 2b); its trace was retrieved in the genome after further analyses (see section below), representing a sequence annotated as a pseudogene. Several teleost species retain two or more ancestral MT1 copies (PP: 0.99, Fig. 1), which can be explained by Ts3R. These duplicates have been annotated in zebrafish as “Mtnr1aa” and “Mtnr1ab” (ZFIN 2008 - ZNC nomenclature, cloned as “ZMel1a1” and “ZMel1a3” by Shang & Zhdanova 2007). Consequently, we maintained the same ‘a’ and ‘b’ nomenclature in all species

according to inferences of orthology with zebrafish (Fig. 1). The two teleost-specific MT1 paralogues were not present in all teleost lineages, with MT1b absent from the studied acanthopterygians (tilapia and pufferfish). Salmonid-specific paralogues of MT1a (MT1a1 and 1a2) were identified, ancestral to three salmonid species (PP: 1.0, Fig. 2b), consistent with retention from Ss4R, though only a single copy of MT1b was retained in the same three species, suggesting ancestral loss following Ss4R (Fig. 1, Fig. 2b).

We identified one copy of MT2 in non-teleost vertebrate lineages, and evidence for teleost-specific paralogues (Fig. 2c). Two MT2 paralogues were identified in Ostariophysi members (zebrafish and Mexican cavefish) and northern pike (Protacanthopterygii); however, only one MT2 copy was identified in Acanthopterygii members (Nile tilapia and pufferfish) (Fig. 1, Fig. 2c). Branching patterns among these duplicates were not well resolved when considering species phylogeny. An ancestral teleost duplication event (e.g. Ts3R) predicts two paralogous MT2 teleost clades, each containing teleosts branching after expected species relationships (as seen for MT1a/b). However, a clade containing zebrafish “Mtnr1ba” (ZFIN 2008, “ZMe1b2” in Shang & Zhdanova 2007) branched outside other fish (including the non-teleost spotted gar) in both the BY and ML trees (Fig. 1 and S1). Internal to the spotted gar, there were two teleost MT2 clades, one containing zebrafish “Mtnr1bb” (ZFIN 2008, “ZMe1b1” in Shang & Zhdanova 2007) and other teleost lineages (northern pike and Acanthopterygii members), while the other contained a separate northern pike sequence and all MT2 sequences from salmonids. Given the strong support for the clade containing zebrafish “Mtnr1bb” (PP:1.0, Bootstrap support: 100%), we considered all sequences therein to be orthologous, and named them MT2b (to maintain the zebrafish “b” nomenclature) (Fig. 2c). We named the

remaining teleost MT2 sequences as MT2a (Fig. 2c), under the hypothesis that orthology to zebrafish MT2a was obscured by a long-branch attraction artefact (note the long-branch length leading to Ostariophysi members for MT2a; Fig. S1). This scenario is parsimonious, as it allows for a single ancestral teleost duplication (e.g. Ts3R) rather than several lineage-specific MT2 gains. Accordingly, we propose that MT2a was lost in the ancestor to *Oreochromis* and *Takifugu*, while two salmonid duplicates of MT2a (MT2a1 and 2a2) were retained from Ss4R (Fig. 1 and S1, Fig. 2c). No copies of MT2b were identified in salmonids, suggesting a loss in the common salmonid ancestor (Fig. 2c).

As shown elsewhere (Dufourny *et al.* 2008), Mel1c and mammalian GPR50 proteins grouped together in a well-supported clade (Fig. 1). A single Mel1c copy was identified in all teleosts barring salmonids, which evidently lack Mel1c (Fig. 2d). This is consistent with a scenario where one Mel1c paralogue was lost early following Ts3R, and an additional loss occurred in the common salmonid ancestor (Fig. 2d).

Synteny analysis supports phylogenetic assignment of vertebrate MTRs

Next, to gain an independent line of evidence to support our phylogenetic reconstructions, we compared the genomic regions harboring MTR-encoding genes among a range of vertebrate lineages. The local gene neighborhood containing each MTR family member was more or less conserved across jawed vertebrate evolution, defining identifiable synteny groups specific to each ancestral MTR (Fig. 3), including teleost and salmonid-specific paralogues (Fig. 4). The genomic neighborhood containing the single *MTR* locus of lampreys did not conserve synteny with an equivalent region containing any single MTR gene in gnathostomes. Instead, the genes surrounding the single *MTR* locus of lampreys showed notable similarity to a

combination of genes located around the various gnathostome MTRs (Fig. 4e). This lends support to an ancestral origin of MTRs in the vertebrate lineage, but does not allow us to pinpoint the relationship of lamprey MTR to the four MTR family members of jawed vertebrates. One possible interpretation is that the duplications generating four gnathostome MTR genes occurred after the cyclostomes and gnathostomes split, with the lamprey genomic neighborhood reflecting a derived representation of the ancestral vertebrate state. However, the current consensus is that at least one round of WGD is shared by cyclostomes and gnathostomes (e.g. Kuraku *et al.* 2009, Stadler *et al.* 2004). In this scenario, conserved synteny between a single genomic region in the former to multiple blocks in the latter may be explained by one or more shared duplications followed by lineage-specific rediploidization, as proposed by Robertson *et al.* 2017.

Genetic linkage between *MTR* and *FAT* genes

Tandem-linked MTR and *FAT* protocadherin gene family members are strongly conserved in all vertebrates (Fig. 3, Fig. 4). Specifically, *MT1*, *Mel1d*, and *MT2* were almost always in tandem with *FAT1*, 2, and 3, respectively (Fig. 3, Fig. 4). This association was absent for *Mel1c*, in addition to MTR co-orthologues from a sea squirt (Fig. 3f) and the Florida lancelet (not shown), defining this as a vertebrate-specific feature. Past studies have noted genetic linkage between MTR and *FAT* genes. For example, the *FAT3*-*MT2* locus is involved in diabetes risk, with several SNPs involved in disease located between the two genes, implying potential functional links (e.g. Prokopenko *et al.* 2009, Dupuis *et al.* 2010). While, the reason for co-evolution of these loci is yet to be determined, the tandem organization of *FAT* and MTR genes indicates selective pressure to maintain an association that may be underpinned by a conserved feature of vertebrate physiology.

308

309 FAT family sequences also provide an independent source of phylogenetic
310 information that may help reconstruct the evolution of the genomic regions containing
311 linked MTR genes. In an ML analysis performed with FAT proteins from
312 representative vertebrate species, we observed three clades (FAT1, 2 and 3) that
313 branched according to expected species relationships (Fig. 5). When the ML tree
314 was midpoint rooted, FAT1 (linked to MT1) and FAT2 (linked to Mel1d) were sister
315 groups (Fig. 5), consistent with the sister relationship of MT1 and Mel1d recovered by
316 the MTR phylogeny. Further, the teleost duplications observed for MTR genes were
317 clearly identifiable in the respective tandem FAT genes (Fig. 5). Finally, the well-
318 supported branching of salmonid FAT3a sequences with zebrafish FAT3a (i.e. linked
319 to the *MT2a* gene, Fig 3c) adds weight to the hypothesis that salmonid/pike MT2
320 sequences are orthologous to zebrafish *MT2a* (Fig. 2c).

321

322 **Synteny analyses support MTR losses**

323 The conservation of synteny across vertebrate taxa in genomic regions containing
324 MTR genes provides useful information on MTR genes inferred to be absent in
325 sequence databases. In this respect, we observed that the genomic regions
326 containing *Mel1d* in reptiles, frogs and fishes have matched syntenic regions in the
327 human and chicken genomes (Fig. 3d). Consequently, the regions predicted to
328 contain *Mel1d* in human and chicken have been properly assembled and are
329 otherwise well annotated, consistent with *bone-fide* genetic losses of *Mel1d* in these
330 species. The same approach allowed us to detect a pseudogene likely to be a
331 vestige of *Mel1c* in Atlantic salmon (LOC106568030) (Fig. 3d), and a gene annotated
332 as 'non-coding' bearing similarity with *MT1* (according to BLAST) at the predicted
333 MT1 locus in spotted gar (LOC107077181) (Fig. 3a). Further, a second *FAT2*

paralogue was detected in Atlantic salmon, supporting our previous conclusion of an ancestral loss of one Mel1d copy following Ss4R. Similarly, a second *FAT3* paralogue was detected in *Oreochromis*, non-paired with an *MTR2* gene (Fig. 3c), confirming the loss of *MT2a* in this species.

Comparative sequence analysis of Mel1d with other MTRs

Having established that Mel1d is an ancestral vertebrate MTR, we sought to compare the primary amino acid sequence of this molecule to other MTR family members, hoping to gain clues on its function considering existing literature (Fig. 6).

We first examined the MTR transmembrane domains and ligand-binding residues, which have known functional importance. The characteristic seven transmembrane domain structure (TMDs) of all MTRs, critical for GPCR structure and ligand binding (Baldwin 1994), were conserved in Mel1d, MT1, MT2 and Mel1c (Fig. 6). Indeed, most of the residues identified as key for melatonin binding are readily identifiable in the Mel1d transmembrane domains (Fig. 6), in particular TM3, 6 and 7 (Gubitz & Reppert 2000, Kokkola *et al.* 2003, Mazna *et al.* 2005, 2008, Chan & Wong 2013). The only notable difference in the TMDs was that several Mel1d orthologues had threonine replacements at position 254, specific to this MTR. This position is important for melatonin binding in MT2 (valine-291 on human MT2), which was not reported for MT1 (Mazna *et al.* 2005). Outside the TMDs, two additional melatonin binding residues (asparagine-102 of the conserved NRY motif and alanine-238) were conserved in Mel1d (Fig. 6). Interestingly, a mutation in the second extracellular loop of GPR50 linked to the loss of melatonin binding function in mammals (Clément *et al.* 2017) was absent in Mel1d (Fig. 6).

Other key sites conserved in Mel1d included cysteine-78 and cysteine-155, responsible for a conserved disulfide bridge essential to MTR structure (Fig. 6). In addition, residues important for G protein activation and trafficking of MT1 (Kokkola *et al.* 2005) were all conserved in Mel1d (green arrows on Fig. 6). Putative palmitoylation site in MT1 and MT2 (cysteine-314 in MT1 and cysteine- 332 in MT2, Sethi *et al.* 2008) required for G-protein interaction (light blue arrow on Fig. 6) were either not identified (MT2 cysteine-332) in Mel1d or absent from most species (MT1 cysteine-314). However a proximal conserved cysteine in position 294 of Mel1d (Fig. 6) may fulfil a similar function. Several phosphorylation sites have been suggested in the C-terminal cytoplasmic tail of MT1 and MT2, which might be important for β -arrestin-dependent receptor internalization (Ebisawa *et al.* 1994, Sethi *et al.* 2008, yellow arrows on Fig. 6). One of these sites is present on Mel1d, at position 288, however only in coelacanth and tetrapods. None of the other phosphorylation sites are present because of the shorter length of Mel1d, and this could be linked to differences in phosphorylation properties.

Residue changes distinguishing Mel1d from other MTRs

The above analyses confirm that Mel1d has most of the canonical residues for melatonin binding and MTR structure/function. We next sought to identify conserved differences between Mel1d and the other MTRs, as candidates to impart functional properties unique to Mel1d.

Five extracellular or intracellular positions in Mel1d show substantial differences with either one or all other MTRs (Fig. 6). In most Mel1d orthologues, the first extracellular loop contains lysine (positive charged) at position 38, which is typically asparagine

(neutrally charged) in the other MTRs. At position 144, which is almost always fixed as glycine in MT1, MT2 and Mel1c, Mel1d orthologues retain glutamic acid or aspartic acid. This replacement is presumed functionally significant, as glycine provides high conformational flexibility (Betts and Russell 2003), while glutamic and aspartic acid are highly negatively charged. At position 246, MTRs usually conserve proline (except for the two derived GPR50 from mammals), but Mel1d shows a high diversity of residues with diverse functional properties, suggesting a distinct mode of selective pressure. In the same loop (position 242), a gap is observed in all Mel1d sequences at an amino acid position that is variable among the other MTRs. Finally, a notable difference between Mel1d and MT1 is observed in position 119, in the second intracellular loop. Most MT1 sequences have aspartic acid at this position, while Mel1d conserves asparagine or serine, leading to a major difference in charge.

DISCUSSION

Our unequivocal demonstration of a new ancestral vertebrate MTR forces a revision of current models for the origin and diversity of MTRs, and has biological implications for vertebrate lineages conserving distinct MTR gene repertoires.

It seems important to ask why Mel1d has previously been missed as a unique MTR, when the gene is readily detectable in sequence databases. This is likely partly due to a historic assumption that the MTR gene family structure of birds and mammals (i.e. MT1, MT2 and Mel1c) is representative for all vertebrates. Mel1d has high similarity with MT1, and has tended to be named 'mtnr1a-like' in genome databases. In addition, previous phylogenetic studies of MTRs have been based on small datasets (e.g. Reppert *et al.* 1995a; Mazurais *et al.* 1999; Park *et al.* 2006, 2007a,b; Shang & Zhdanova 2007; Hong *et al.* 2014), with biases in the taxa sampled, and could not by design distinguish Mel1d and MT1. A single past study noted a *Xenopus* MTR sequence that did not group with MT1, MT2 or Mel1c and concluded the existence of a novel MTR (Mel1d) (Shiu *et al.* 1996); correctly according to our findings. Our study benefits from a much broader survey of vertebrate MTR sequences, allowing us to conclude that Mel1d is at least 450 million years old, having been present in the jawed vertebrate ancestor.

Our phylogenetic reconstruction of MTRs will help the field going forwards, as researchers can be certain of which family member (including teleost-specific paralogues) they are studying, allowing more reliable conclusions in comparative studies of function and gene expression. We show that teleost-specific paralogues of MT1 are easily distinguished from Mel1d and provide a scheme to allow researchers

to match teleost MTRs formerly named under several nomenclature systems to a single phylogenetically-assigned naming system accommodating orthologues and paralogues (Table 1).

Insights into Mel1d function: reinterpreting expression data in teleosts

While not being previously recognized as a unique vertebrate MTR, Mel1d has already been studied in various teleosts (Table 1). These past studies demonstrate that the *Mel1d* transcript is abundantly expressed in a manner like other MTR family members, but showing differences that may underlie unique functions. A pattern seems conserved across multiple species, where Mel1d and MT1a expression is higher in brain and retina, respectively (e.g. Park *et al.* 2006, 2007a,b; Ikegami *et al.* 2009; Confente *et al.* 2010; Hong *et al.* 2014). Mel1d tends to be more strongly expressed in brain regions associated with visual perception (e.g. Mazurais *et al.* 1999; Gaildrat and Falcón, 2000; Shi *et al.* 2004; Confente *et al.* 2010; Hong *et al.* 2014). Many peripheral tissues were reported to express Mel1d with species-specific differences and in a distinct manner to other MTRs (Park *et al.* 2006, 2007a,b; Ikegami *et al.* 2009; Confente *et al.* 2010; Hong *et al.* 2014). Such data suggests involvement of Mel1d in photoreceptive processes, along with broader regulatory roles in the physiological functions of peripheral organs.

Rhythmical oscillations in the expression of Mel1d have also been reported, with variations depending on species, organ and season. In zebrafish, a day/night oscillation of MTR brain gene expression (peaking at night) was noted for all six MTR paralogues, including Mel1d, with further expression upregulation in response to melatonin administration (Shang & Zhdanova 2007). In golden rabbitfish, MT1a, Mel1d and Mel1c expression was higher at night for brain and retina, with Mel1d

levels peaking at different times (Park *et al.* 2006, 2007a,b, 2014). In goldfish, Mel1d was the only MTR showing rhythmical oscillations in optic tectum expression, while the same was true for MT1a in retina, both peaking at the night-day transition (Ikegami *et al.* 2008). In a marine pufferfish, Mel1d, MT1a and Mel1c showed synchronous daily cycling of expression in the pineal gland with a nocturnal peak (Ikegami *et al.* 2015). Conversely, in golden rabbitfish pineal gland, oscillations were desynchronized for the same three MTRs (Park *et al.* 2006, 2007a,b). Daily rhythmicity in Mel1d expression has also been observed in peripheral tissues (liver and kidney) of golden rabbitfish, with higher expression during the day, opposite to the brain/retina (Park *et al.* 2006, 2007b). In addition to daily variation in regulation, Mel1d expression is regulated by other cycles, for example showing semilunar oscillation in the diencephalon of mudskipper (Hong *et al.* 2014) and ultradiurnal oscillation in a marine pufferfish, which may be circatidal (Ikegami *et al.* 2015). Mel1d expression in the Senegalese sole exhibited stronger day-to-night and seasonal variation than other MTR family members, with reciprocal differences recorded between retina and optic tectum (Confente *et al.* 2010). Therefore, past work shows that Mel1d is regulated during multiple biological cycles in teleosts, showing variations distinct from other MTRs, implying functional distinctiveness.

Functional divergence between Mel1d and MT1?

High protein-level similarity between Mel1d and MT1, taken with the conservation of all key residues in the MTR transmembrane domains, strongly implies that Mel1d binds melatonin. Notably, residues showing conserved replacements between Mel1d and MT1 are all located in extracellular or cytoplasmic loops, which is predicted to impact interactions with other proteins, in particular signalling partners, rather than melatonin. Strikingly, one of these sites corresponds to a documented human MT1

mutation studied *in vitro* (Chaste *et al.* 2010). The replacement of glycine-144 (MT1) with glutamic acid or aspartic acid corresponds to a G166E mutation in human MT1, associated with impaired activation of cAMP signalling, despite retention of strong melatonin binding (Chaste *et al.* 2010). The elephant shark retains glutamic acid at this position in both MT1 and Mel1d, suggesting this represents the ancestral state, with functional divergence arising in the common ancestor to lobe and ray-finned fishes. It is also intriguing to observe that Mel1d of two tetrapods have apparently reverted to glycine in this position, indicating selection towards the ancestral residue.

Why was Mel1d lost in mammals and birds?

Further work will be needed to establish the extent of conservation in Mel1d function and regulation across different vertebrate lineages. This should focus on reptiles and amphibians, where the function of this gene has not been studied experimentally. Such studies may help explain the specific biological requirements for Mel1d, and reveal why the gene was lost independently in mammals and birds. It is notable that mammals and birds stand out from other vertebrates when considering their melatonin-dependent light detection and clock systems. Mammals have lost extraocular light perception and relocated control of their biological clock away from melatonin-producing pinealocytes to the suprachiasmatic nucleus (Falcón *et al.* 2009). Birds have both the ancestral pineal clock and melatonin production system, but also independently developed a clock system in the homologue of the suprachiasmatic nucleus and use retinal detection (Cassone 1991, Falcón *et al.* 2009). Another distinguishing feature specific to both groups is homeothermy, with modulatory effects of melatonin on body temperature regulation reported in humans (Cagnacci *et al.* 1992; Viswanathan *et al.* 1990) and Japanese quail (Underwood and Edmonds 1995). Extrinsic temperature variation appears a less important zeitgeber

for the circadian clock of homeotherms relative to poikilotherms (Rensing and Ruoff, 2002), which are known to use melatonin to regulate behavioral thermoregulation (Lutterschmidt *et al.* 2003). In addition, birds and mammals are the only vertebrates that have evolved (through convergent mechanisms) stereotypical slow wave and rapid eye movement sleep phases, linked to melatonin regulation in mammals (Lesku *et al.* 2011). Such changes in the physiological role of melatonin and consequent re-organization of melatonin response pathways, may have been the ultimate driver for Mel1d redundancy and gene loss through relaxation of purifying selection.

Another melatonin-associated function that is present in vertebrate lineages retaining Mel1d (in addition to lamprey), but lost in both mammals and birds, is the negative regulation of pigmentation development in the dark, known as the “body-blanching response” (Hamasaki and Eder 1977, Norris and Carr 2013). In fishes, melatonin is thought to regulate chromatosome aggregation in different kinds of chromatophores (Fujii 2000); Mel1d is expressed in the skin of mudskipper (together with MT1 - Hong *et al.* 2014), the goldfish (together with MT2 and Mel1c - Ikegami *et al.* 2008) and the sole (together with MT2 - Confente *et al.* 2010). In addition, in sole skin, Mel1d is the only MTR to be up-regulated at night. It is therefore possible that Mel1d is involved in skin physiology and pigment regulation in fish chromatophores.

Expansion of the MTR repertoire of teleosts

Contrary to mammals/birds, there has been a trend towards evolutionary expansion in the MTR repertoire of teleosts, as observed in many gene families with paralogues retained from Ts3R (Glasauer and Neuhauss, 2014) and Ss4R (Houston and Macqueen, 2019). Interestingly, not all MTR family members were affected equally. While we identified multiple paralogous copies of MT1 and MT2 - presumed to have

been retained from Ts3R and Ss4R - Mel1c and Mel1d were always single copy, requiring repeated losses of paralogues generated during gene duplication or WGD events. This is compatible with a hypothesis where the functions or expression-level regulation of MT1 and MT2 can be divided among paralogous copies, following the well-established subfunctionalization model, or potentially reflects fixation of new adaptive functions among MT1/MT2 paralogues (Stoltzfus 1999 and Force *et al.* 1999). In this respect, we observed several amino acid substitutions between MT1a vs. MT1b and MT2a vs. MT2b (Fig. 6), consistent with protein-level functional divergence. Conversely, selection has operated in a distinct manner for Mel1c and Mel1d, with any duplicates generated being quickly purged by selection for reasons that remain to be established, but potentially linked to dosage constraints, or a mechanism of regulation that cannot be divided across distinct loci.

CONCLUSIONS

Mel1d is one of four ancestral vertebrate MTRs that shows a wide phylogenetic distribution but has been lost in mammals and birds. Compared to MT1, MT2 and Mel1c, Mel1d has many conserved, but also divergent characteristics, both in terms of protein sequence and spatio-temporal expression patterns of relevance to chronobiological traits. Additional work is needed to characterize the functional distinctiveness of Mel1d compared to other MTRs and to explain why unique MTR repertoires have been conserved in different vertebrate lineages.

564

565 Conflict of interest: The authors declare no conflict of interest.

566

567 Acknowledgments: This work was funded by the Frimmedbio (Fri prosjektstøtte for
568 medisin, helse og biologi) program of the Research Council of Norway (grant number
569 241016 - "Light & Salt - Thyroid hormone deiodinase paralogues & the evolution of
570 complex life-history strategy in salmonids"). DJM was supported from BBSRC
571 Institute Strategic Programme funding to The Roslin Institute (grant ref:
572 BBS/E/D/10002071). The study was initiated by DH, LOEE and DJM. Sequence
573 collection, alignment and phylogenetic analyses was done by DJM. Comparative
574 genomic and sequence analyses were done by ED. The manuscript was written by
575 ED and DJM, with critical inputs from DH and LOEE.

REFERENCES

- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17): 3389-3402.
- Baldwin, J.M., 1994 Structure and function of receptors coupled to G proteins. *Curr. Opin. Cell Biol.* 6: 180-190.
- Betts, M.J., Russell, R.B., 2003 Amino acid properties and consequences of substitutions, in *Bioinformatics for Geneticists*, M.R. Barnes, I.C. Gray eds, Wiley.
- Cagnacci, A., Elliott, J.A., Yen, S.S., 1992 Melatonin: a major regulator of the circadian rhythm of core temperature in humans. *J Clin Endocrinol Metab.* 75: 447-452.
- Cassone, V.M., 1991 Melatonin and suprachiasmatic nucleus function, pp. 309–323 in *Suprachiasmatic Nucleus: The Mind's Clock*, D.C. Klein, R.Y. Moore, S.M. Reppert, Oxford Univ. Press, New York.
- Chan, K.H., Wong, Y.H., 2013 A molecular and chemical perspective in defining MTR subtype selectivity. *Int J Mol Sci.* 14: 18385-18406.
- Chai, K., Liu, X., Zhang, Y., Lin. H., 2013 Day-night and reproductive cycle profiles of melatonin receptor, kiss, and gnhrh expression in orange-spotted grouper (*Epinephelus coioides*). *Mol Reprod Dev.* 80: 535-548.

602

603 Chaste, P., Clément, N., Mercati, O., Guillaume, J.L., Delorme, R., Botros, H.G., *et*
604 *al.*, 2010 Identification of pathway-biased and deleterious MTR mutants in autism
605 spectrum disorders and in the general population. PLoS One. 5: e11495.

606

607 Chiari, Y., Cahais, V., Galtier, N., Delsuc, F., 2012 Phylogenomic analyses support
608 the position of turtles as the sister group of birds and crocodiles (Archosauria). BMC
609 Biol. 10: 65.

610

611 Choi, Y.J., Habibi, H.R., Choi, C.Y., 2016 Profiles of gonadotropin-inhibitory hormone
612 and melatonin during the sex change and maturation of cinnamon clownfish,
613 *Amphiprion melanopus*. Biochem Biophys Res Commun. 475: 189-93.

614

615 Clément, N., Renault, N., Guillaume, J.L., Cecon, E., Journé, A.S., Laurent, X., *et al.*
616 2018 Importance of the second extracellular loop for melatonin MT1 receptor function
617 and absence of melatonin binding in GPR50. Br J Pharmacol. 175: 3281-3297.

618

619 Confente, F., Rendón, M., Besseau, L., Falcón, J., Muñoz-Cueto, J.A., 2010 MTRs in
620 a pleuronectiform species, *Solea senegalensis*: Cloning, tissue expression, day-night
621 and seasonal variations. Gen Comp Endocrinol. 167: 202-214.

622

623 Dehal, P., Boore, J.L., 2005 Two rounds of whole genome duplication in the
624 ancestral vertebrate. PLoS Biol. 3: e314.

625

626 Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006 Relaxed
627 phylogenetics and dating with confidence. PLoS Biol. 4: e88.

628

629 Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012 Bayesian phylogenetics
630 with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29: 1969-1973.

631

632 Dubocovich, M.L., Markowska, M., 2005 Functional MT1 and MT2 MTRs in
633 Mammals. Endocrine 27: 101-110.

634

635 Dufourny, L., Levasseur, A., Migaud, M., Callebaut, I., Pontarotti, P., Malpoux, B.,
636 Monget, P., 2008 GPR50 is the mammalian ortholog of Mel1c: Evidence of rapid
637 evolution in mammals. BMC Evol. Biol. 8: 105.

638

639 Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U.
640 *et al.*, 2010 New genetic loci implicated in fasting glucose homeostasis and their
641 impact on type 2 diabetes risk. Nat Genet. 42: 105-116.

642

643 Ebisawa, T., Karne, S., Lerner, M.R., Reppert, S.M., 1994 Expression cloning of a
644 high-affinity MTR from *Xenopus* dermal melanophores. Proc Natl Acad Sci USA. 91:
645 6133-6137.

646

647 Falcón, J., Besseau, L., Fuentès, M., Sauzet, S., Magnanou, E., Boeuf, G., 2009
648 Structural and functional evolution of the pineal melatonin system in vertebrates. Ann
649 N Y Acad Sci. 1163: 101-111.

650

651 Feng, N.Y., Bass, A.H., 2016 "Singing" Fish Rely on Circadian Rhythm and
652 Melatonin for the Timing of Nocturnal Courtship Vocalization. Curr Biol. 26: 2681-
653 2689.

654

655 Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., Postlethwait, J., 1999
656 Preservation of duplicate genes by complementary, degenerative mutations.
657 Genetics. 151: 1531-1545.

658

659 Fujii, R., 2000 The regulation of motile activity in fish chromatophores. Pigment Cell
660 Res. 13: 300-319.

661

662 Gaildrat, P., Falcón, J., 2000 Melatonin receptors in the pituitary of a teleost fish:
663 mRNA expression, 2-[(125)I]iodomelatonin binding and cyclic AMP response.
664 Neuroendocrinology. 72: 57-66.

665

666 Gauer, F., Masson-Pévet, M., Stehle, J., Pevet, P., 1994 Daily variations in MTR
667 density of rat pars tuberalis and suprachiasmatic nuclei are distinctly regulated. Brain
668 Res 641: 92-98.

669

670 Gernhard, T., 2008 The conditioned reconstructed process. J Theor Biol. 253: 769-
671 778.

672

673 Glasauer, S.M., Neuhauss, S.C., 2014 Whole-genome duplication in teleost fishes
674 and its evolutionary consequences. Mol Genet Genomics. 289: 1045-1060.

675

676 Gubit, A.K., Reppert, S.M., 2000 Chimeric and point-mutated receptors reveal that a
677 single glycine residue in transmembrane domain 6 is critical for high affinity
678 melatonin binding. Endocrinology 141: 1236-1244.

679

680 Hardeland, R., Cardinali, D.P., Srinivasan, V., Spence, D.W., Brown, G.M., Pandi-
 681 Perumal, S.R., 2011 Melatonin--a pleiotropic, orchestrating regulator molecule. *Prog*
 682 *Neurobiol.* 93: 350-384.
 683
 684 Hamasaki, D.I., Eder, D.J., 1977 Adaptive Radiation of the Pineal System. pp.
 685 7(5)497–548 in *Handbook of sensory physiology - The Visual System in Vertebrates*
 686 (F. Crescitelli, ed.), Springer-Verlag, New York.
 687
 688 Herrera-Pérez, P., Del Carmen Rendón, M., Besseau, L., Sauzet, S., Falcón, J.,
 689 Muñoz-Cueto, J.A., 2010 MTRs in the brain of the European sea bass: An *in situ*
 690 hybridization and autoradiographic study. *J Comp Neurol.* 518: 3495-3511.
 691
 692 Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018
 693 UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35: 518-
 694 522.
 695
 696 Hong, L.Y., Hong, W.S., Zhu, W.B., Shi, Q., You, X.X., Chen, S.X., 2014 Cloning and
 697 expression of MTRs in the mudskipper *Boleophthalmus pectinirostris*: their role in
 698 synchronizing its semilunar spawning rhythm. *Gen Comp Endocrinol.* 195: 138-150.
 699
 700 Houston, R.D., Macqueen, D.J., 2019 Atlantic salmon (*Salmo salar* L.) genetics in the
 701 21st century: taking leaps forward in aquaculture and biological understanding. *Anim*
 702 *Genet.* 50: 3-14.
 703

704 Ikegami, T., Azuma, K., Nakamura, M., Suzuki, N., Hattori, A., Ando, H., 2009 Diurnal
 705 expressions of four subtypes of MTR genes in the optic tectum and retina of goldfish.
 706 Comp Biochem Physiol A Mol Integr Physiol. 152: 219-224.
 707

708 Ikegami, T., Maruyama, Y., Doi, H., Hattori, A., Ando, H., 2015 Ultradian oscillation in
 709 expression of four MTR subtype genes in the pineal gland of the grass puffer, a
 710 semilunar-synchronized spawner, under constant darkness. Front Neurosci. 9: 9.
 711

712 Jin, Y.H., Park, J.W., Kim, J.H., Kwon, J.Y., 2013 The Expression Pattern of MTR 1a
 713 Gene during Early Life Stages in the Nile tilapia (*Oreochromis niloticus*). Dev Reprod.
 714 17: 45-53.
 715

716 Jockers, R., Maurice, P., Boutin, J.A., Delagrange, P., 2008 MTRs,
 717 heterodimerization, signal transduction and binding sites: what's new? Br J
 718 Pharmacol. 154: 1182-1195.
 719

720 Jones, D.T., Taylor, W.R., Thornton, J.M., 1992 The rapid generation of mutation
 721 data matrices from protein sequences. Bioinformatics 8: 275-282.
 722

723 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S.,
 724 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat
 725 Methods. 14: 587-589.
 726

727 Kamesh, N., Aradhyam, G.K., Manoj, N., 2008 The repertoire of G protein-coupled
 728 receptors in the sea squirt *Ciona intestinalis*. BMC Evol. Biol. 8: 129.
 729

730 Katoh, K., Standley, D.M., 2013 MAFFT multiple sequence alignment software
 731 version 7: improvements in performance and usability. Mol Biol Evol. 30: 772-780.
 732

733 Kokkola, T., Foord, S.M., Watson, M.A., Vakkuri, O., Laitinen, J.T., 2003 Important
 734 amino acids for the function of the human MT1 MTR. Biochem Pharmacol. 65: 1463-
 735 1471.
 736

737 Kokkola, T., Salo, O.M., Poso, A., Laitinen, J.T., 2005 The functional role of cysteines
 738 adjacent to the NRY motif of the human MT1 MTR. J Pineal Res. 39: 1-11.
 739

740 Krishnan, A., Almén, M.S., Fredriksson, R., Schiöth, H.B., 2013 Remarkable
 741 similarities between the hemichordate (*Saccoglossus kowalevskii*) and vertebrate
 742 GPCR repertoire. Gene. 526:122-133.
 743

744 Kuraku, S., 2013 Impact of asymmetric gene repertoire between cyclostomes and
 745 gnathostomes. Semin. Cell Dev. Biol. 24: 119-127.
 746

747 Kuraku, S., Meyer, A., Kuratani, S., 2009 Timing of Genome Duplications Relative to
 748 the Origin of the Vertebrates: Did Cyclostomes Diverge before or after? Mol Biol Evol
 749 26: 47-59.
 750

751 Lesku, J.A., Vyssotski, A.L., Martinez-Gonzalez, D., Wilzeck, C., Rattenborg, N.C.,
 752 2011 Local sleep homeostasis in the avian brain: convergence of sleep function in
 753 mammals and birds? Proc. R. Soc. B 278: 2419-2428.
 754

755 Levoye, A., Dam, J., Ayoub, M.A., Guillaume, J.L., Couturier, C., Delagrangé, P.,
 756 Jockers, R., 2006 The orphan GPR50 receptor specifically inhibits MT1 MTR function
 757 through heterodimerization. *Embo J.* 25: 3012-3023.
 758
 759 Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., *et al.*, 2016
 760 The Atlantic salmon genome provides insights into rediploidization. *Nature.* 533: 200-
 761 205.
 762
 763 Lutterschmidt, D.I., Lutterschmidt, W.I., Hutchison, V.H., 2003 Melatonin and
 764 thermoregulation in ectothermic vertebrates: a review. *Can. J. Zool.* 81: 1-13.
 765
 766 Macqueen, D.J., Johnston, I.A., 2014 A well-constrained estimate for the timing of
 767 the salmonid whole genome duplication reveals major decoupling from species
 768 diversification. *Proc Biol Sci.* 281: 20132881.
 769
 770 Macqueen, D.J., Wilcox, A.H., 2014 Characterization of the definitive classical
 771 calpain family of vertebrates using phylogenetic, evolutionary and expression
 772 analyses. *Open Biol.* 4: 130219.
 773
 774 Mazna, P., Grycova, L., Balik, A., Zemkova, H., Friedlova, E., Obsilova, V., *et al.*,
 775 2008 The role of proline residues in the structure and function of human MT2 MTR. *J*
 776 *Pineal Res.* 45: 361-372.
 777
 778 Mazna, P., Berka, K., Jelinkova, I., Balik, A., Svoboda, P., Obsilova, V., *et al.*, 2005
 779 Ligand binding to the human MT2 MTR: the role of residues in transmembrane
 780 domains 3, 6, and 7. *Biochem Biophys Res Commun.* 332: 726-734.

781

782 Mazurais, D., Brierley, I., Anglade, I., Drew, J., Randall, C., Bromage, N., *et al.*, 1999

783 Central MTRs in the rainbow trout: comparative distribution of ligand binding and

784 gene expression. *J. Comp. Neurol.* 409: 313-324.

785

786 Norris, D.O., Carr, J.A., 2013 *Vertebrate endocrinology*. Academic Press.

787

788 Nordstrom, K.J., Fredriksson, R., Schioth, H.B., 2008 The amphioxus *Branchiostoma*

789 *floridae* genome contains a highly diversified set of G protein-coupled receptors.

790 *BMC Evol. Biol.* 8: 9.

791

792 Nguyen, L-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015 IQ-TREE: A fast

793 and effective stochastic algorithm for estimating maximum likelihood

794 phylogenies. *Mol. Biol. Evol.* 32: 268-274.

795

796 Nguyen, N.T.T., Vincens, P., Roest Crollius, H., Louis, A., 2018 Genomicus 2018:

797 karyotype evolutionary trees and on-the-fly synteny computing. *Nucleic Acids Res.*

798 46: D816-D822.

799

800 Ota, K.G., Fujimoto, S., Oisi, Y., Kuratani, S., 2011 Identification of vertebra-like

801 elements and their possible differentiation from sclerotomes in the hagfish. *Nat*

802 *Commun.* 2: 373.

803

804 Park, Y.J., Park, J.G., Takeuchi, Y., Hur, S.P., Lee, Y.D., Kim, S.J., Takemura, A.,

805 2014 Influence of moonlight on mRNA expression patterns of MTR subtypes in the

806 pineal organ of a tropical fish. *Mar Genomics.* 14: 67-70.

807

808 Park, Y.J., Park, J.G., Jeong, H.B., Takeuchi, Y., Kim, S.J., Lee, Y.D., Takemura, A.,
809 2007(a) Expression of the MTR Mel(1c) in neural tissues of the reef fish *Siganus*
810 *guttatus*. Comp Biochem Physiol A Mol Integr Physiol. 147: 103-111.

811

812 Park, Y.J., Park, J.G., Hiyakawa, N., Lee, Y.D., Kim, S.J., Takemura, A., 2007(b)
813 Diurnal and circadian regulation of a MTR, MT1, in the golden rabbitfish, *Siganus*
814 *guttatus*. Gen Comp Endocrinol. 150: 253-262.

815

816 Park, Y.J., Park, J.G., Kim, S.J., Lee, Y.D., Saydur Rahman, M., Takemura, A., 2006
817 MTR of a reef fish with lunar-related rhythmicity: cloning and daily variations. J Pineal
818 Res. 41: 166-174.

819

820 Prokopenko, I., Langenberg, C., Florez, J.C., Saxena, R., Soranzo, N., Thorleifsson,
821 G., *et al.*, 2009 Variants in MTNR1B influence fasting glucose levels. Nat. Genet. 41:
822 77-81.

823

824 Redmond, A.K., Macqueen, D.J., Dooley, H., 2018 Phylotranscriptomics suggests
825 the jawed vertebrate ancestor could generate diverse helper and regulatory T cell
826 subsets. BMC Evol Biol. 18: 169.

827

828 Rensing, L., Ruoff, P., 2002 Temperature effect on entrainment, phase shifting, and
829 amplitude of circadian clocks and its molecular bases. Chronobiol Int. 19: 807-864.

830

831 Reppert, S.M., Weaver, D.R., Ebisawa, T., 1994 Cloning and characterization of a
832 mammalian MTR that mediates reproductive and circadian responses. *Neuron* 13:
833 1177-1185.

834

835 Reppert, S.M., Godson, C., Mahle, C.D., Weaver, D.R., Slaugenhaupt, S.A., Gusella,
836 J.F., 1995(a) Molecular characterization of a second MTR expressed in human retina
837 and brain: the Mel1b-MTR. *Proc Natl Acad Sci USA* 92: 8734-8738.

838

839 Reppert, S.M., Weaver, D.R., Cassone, V.M., Godson, C., Kolakowski, L.F. Jr., 1995
840 (b) Melatonin receptors are for the birds: molecular analysis of two receptor subtypes
841 differentially expressed in chick brain. *Neuron* 15: 1003-1015.

842

843 Robertson, F.M., Gundappa, M.K., Grammes, F., Hvidsten, T.R., Redmond, A.K.,
844 Lien, S., *et al.*, 2017 Lineage-specific rediploidization is a mechanism to explain time-
845 lags between genome duplication and evolutionary diversification. *Genome Biol.* 18:
846 111.

847

848 Sauzet, S., Besseau, L., Herrera Perez, P., Covès, D., Chatain, B., Peyric, E., *et al.*,
849 2008 Cloning and retinal expression of MTRs in the European sea bass,
850 *Dicentrarchus labrax*. *Gen Comp Endocrinol.* 157: 186-195.

851

852 Sethi, S., Adams, W., Pollock, J., Witt-Enderby, P.A., 2008, C-terminal domains
853 within human MT1 and MT2 MTRs are involved in internalization processes. *J Pineal*
854 *Res.* 45: 212-218.

855

856 Shang, E.H., Zhdanova, I.V., 2007 The circadian system is a target and modulator of
857 prenatal cocaine effects. PLoS One. 2: e587.

858

859 Shi, Q., Ando, H., Coon, S.L., Sato, S., Ban, M., Urano, A., 2004 Embryonic and
860 post-embryonic expression of arylalkylamine N-acetyltransferase and MTR genes in
861 the eye and brain of chum salmon (*Oncorhynchus keta*). Gen Comp Endocrinol. 136:
862 311-321.

863

864 Shiu, S.Y., Ng, N., Pang, S.F., 1996 A molecular perspective of the genetic
865 relationships of G-protein coupled MTR subtypes. J Pineal Res. 20: 198-204.

866

867 Slominski, R.M., Reiter, R.J., Schlabritz-Loutsevitch, N., Ostrom, R.S., Slominski,
868 A.T., 2012 Melatonin membrane receptors in peripheral tissues: distribution and
869 functions. Mol Cell Endocrinol. 351: 152-166.

870

871 Solovyev, V., Kosarev, P., Seledsov, I., Vorobyev, D., 2006 Automatic annotation of
872 eukaryotic genes, pseudogenes and promoters. Genome Biol. 7: 10.1-10.12.

873

874 Stadler, P.F., Fried, C., Prohaska, S.J., Bailey, W.J., Misof, B.Y., Ruddle, F.H.,
875 Wagner, G.P., 2004 Evidence for independent Hox gene duplications in the hagfish
876 lineage: a PCR-based gene inventory of *Eptatretus stoutii*. Mol Phylogenet Evol.
877 32:686-694.

878

879 Stoltzfus, A., 1999 On the possibility of constructive neutral evolution. J Mol Evol. 49:
880 169-181.

881

882 Sugahara, F., Murakami, Y., Adachi, N., Kuratani, S., 2013 Evolution of the
 883 regionalization and patterning of the vertebrate telencephalon: what can we learn
 884 from cyclostomes? *Curr. Opin. Genet. Dev.* 23: 475-483.

885

886 Talavera, G., Castresana, J., 2007 Improvement of phylogenies after removing
 887 divergent and ambiguously aligned blocks from protein sequence alignments. *Syst*
 888 *Biol.* 56: 564-577.

889

890 Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016 W-IQ-TREE: a
 891 fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44:
 892 W232-235.

893

894 Underwood, H., Edmonds, K., 1995 The circadian rhythm of thermoregulation in
 895 Japanese quail: III. Effects of melatonin administration. *J Biol Rhythms.* 10: 284-298.

896

897 Vernadakis, A.J., Bemis, W.E., Bittman, E.L., 1998 Localization and partial
 898 characterization of MTRs in amphioxus, hagfish, lamprey, and skate. *Gen Comp*
 899 *Endocrinol.* 110: 67-78.

900

901 Vilella, A.J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., Birney, E., 2009
 902 EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in
 903 vertebrates. *Genome Res.* 19: 327-335.

904

905 Viswanathan, M., Laitinen, J.T., Saavedra, J.M., 1990 Expression of MTRs in arteries
 906 involved in thermoregulation. *Proc Natl Acad Sci USA* 87: 6200-6203.

907

908 Weaver, D.R., Reppert, S.M., 1990 Melatonin receptors are present in the ferret pars
909 tuberalis and pars distalis, but not in brain. *Endocrinology*. 127: 2607-2609.
910
911 Witt-Enderby, P.A., Bennett, J., Jarzynka, M.J., Firestine, S., Melan, M.A., 2003
912 MTRs and their regulation: biochemical and structural mechanisms. *Life Sciences* 72:
913 2183-2198.
914
915 ZFIN, <https://zfin.org>

FIGURE LEGENDS

Fig. 1. Bayesian phylogenetic tree of MTR family evolution in jawed vertebrates. The analysis was done using BEAST with a high-confidence alignment of eighty MTRs (300 amino acid positions; Additional Dataset 1), an uncorrelated relaxed molecular clock model and the best-fitting amino acid substitution model (JTT+F+I+G4). Numbers on branches are posterior probability support. Three WGD events in vertebrate evolution are shown (2R - ancestral to vertebrates; Ts3R - ancestral to teleosts; Ss4R - ancestral to salmonids). A ML tree was performed using the same data and is provided in Fig. S1.

Fig. 2. Proposed evolutionary history of each MTR family member, considering (a) Mel1d, (b) MT1, (c) MT2m, and (d) Mel1c. Species inferred to have lost all copies of a MTR gene are highlighted in dark red. Teleost species inferred to have lost paralogues of MTR genes arising from the Ts3R and Ss4R events are highlighted in light red.

Fig. 3. Conserved synteny between the genomic neighbourhood containing MTR orthologues of different lineages, shown for (a) jawed vertebrate MT1, (b) jawed vertebrate Mel1d, (c) jawed vertebrate MT2, (d) jawed vertebrate Mel1c, (e) comparing MTR from two lamprey species with jawed vertebrates, and (f) comparing a urochordate with vertebrates.

Fig. 4. Conserved synteny between the genomic neighbourhood containing MTR paralogues retained from Ts3R and Ss4R, shown for (a) MT1a, (b) MT1b, (c) Mel1d, (d) MT2a, (e) MT2b, and (f) Mel1c.

942

943 Fig. 5. ML phylogenetic analysis of FAT atypical protocadherins in jawed vertebrates.
944 The analysis was done using IQ-TREE with a high-confidence alignment of thirty-five
945 FAT proteins (2,540 amino acid positions; Additional Dataset 2) and the best-fitting
946 amino acid substitution model (JTT+G+I). Numbers on branches are bootstrap
947 support values. Other details are as in the Fig. 1 legend.

948

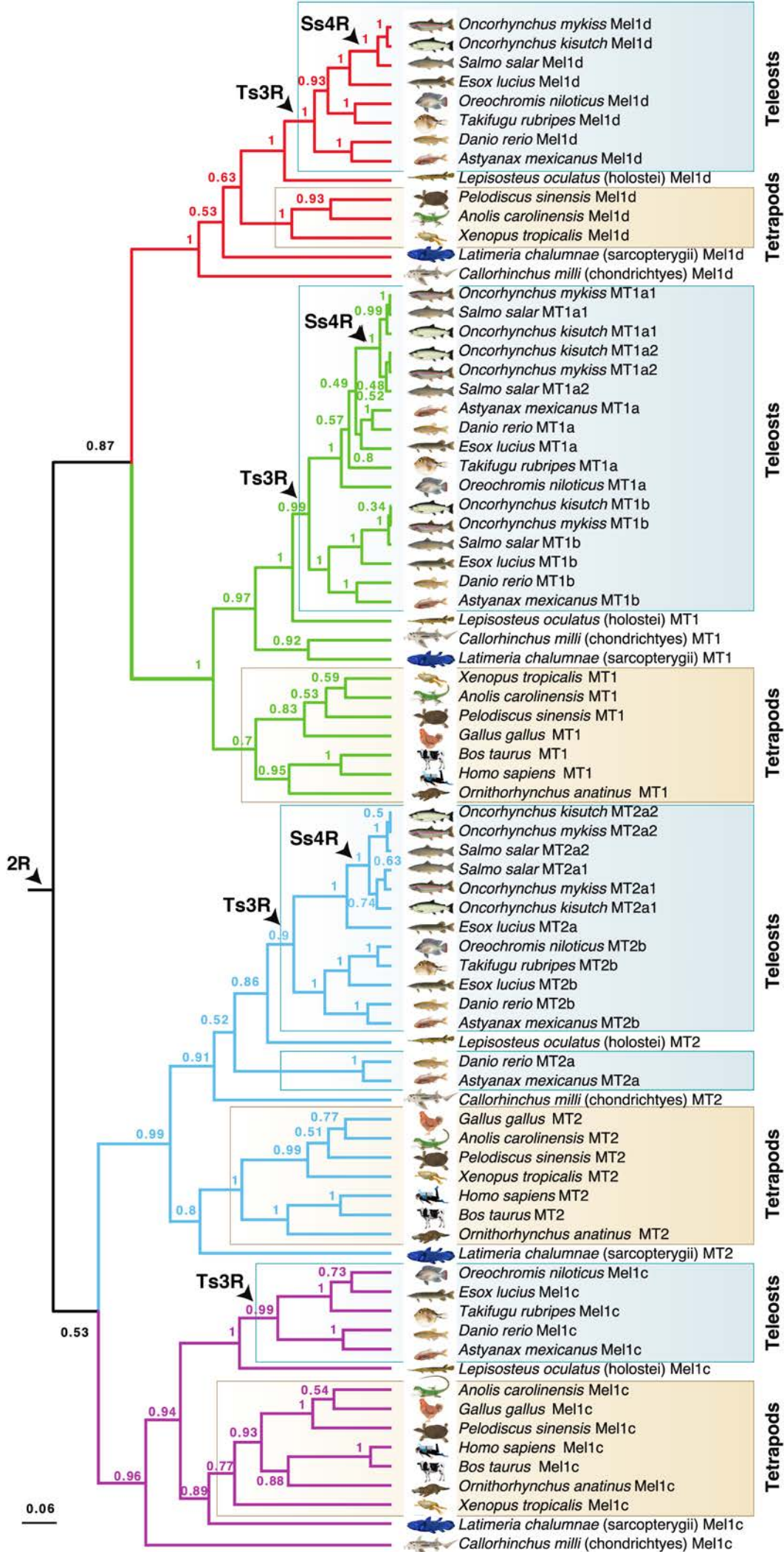
949 Fig. 6. Alignment used to compare amino acid positions among vertebrate MTR
950 proteins (matching to the alignment used for phylogenetic analysis; Additional
951 Dataset S1). Species abbreviations: Ac = *Anolis carolensis* (green anole lizard); Am
952 = *Astyanax mexicanus* (Mexican cavefish); Bt = *Bos taurus* (cattle); Cm =
953 *Callorhinchus milli* (elephant shark); Dr = *Danio rerio* (zebrafish); El = *Esox lucius*
954 (northern pike); Gg = *Gallus gallus* (chicken); Hs = *Homo sapiens* (human); Lc =
955 *Latimeria chalumnae* (coelacanth); Lo = *Lepisosteus oculatus* (spotted gar); Oa =
956 *Ornithorhynchus anatinus* (platypus); On = *Oreochromis niloticus* (Nile tilapia); Ps =
957 *Pelodiscus sinensis* (Chinese softshell turtle); Tr = *Takifugu rubripes* (tiger pufferfish);
958 Xt = *Xenopus tropicalis* (western clawed frog). Detailed annotation of sequences
959 flagged up in the main text are provided within the figure.

960

961 Table 1. Phylogenetic assignment of teleost MTRs to a standardized nomenclature
962 system.

963

964

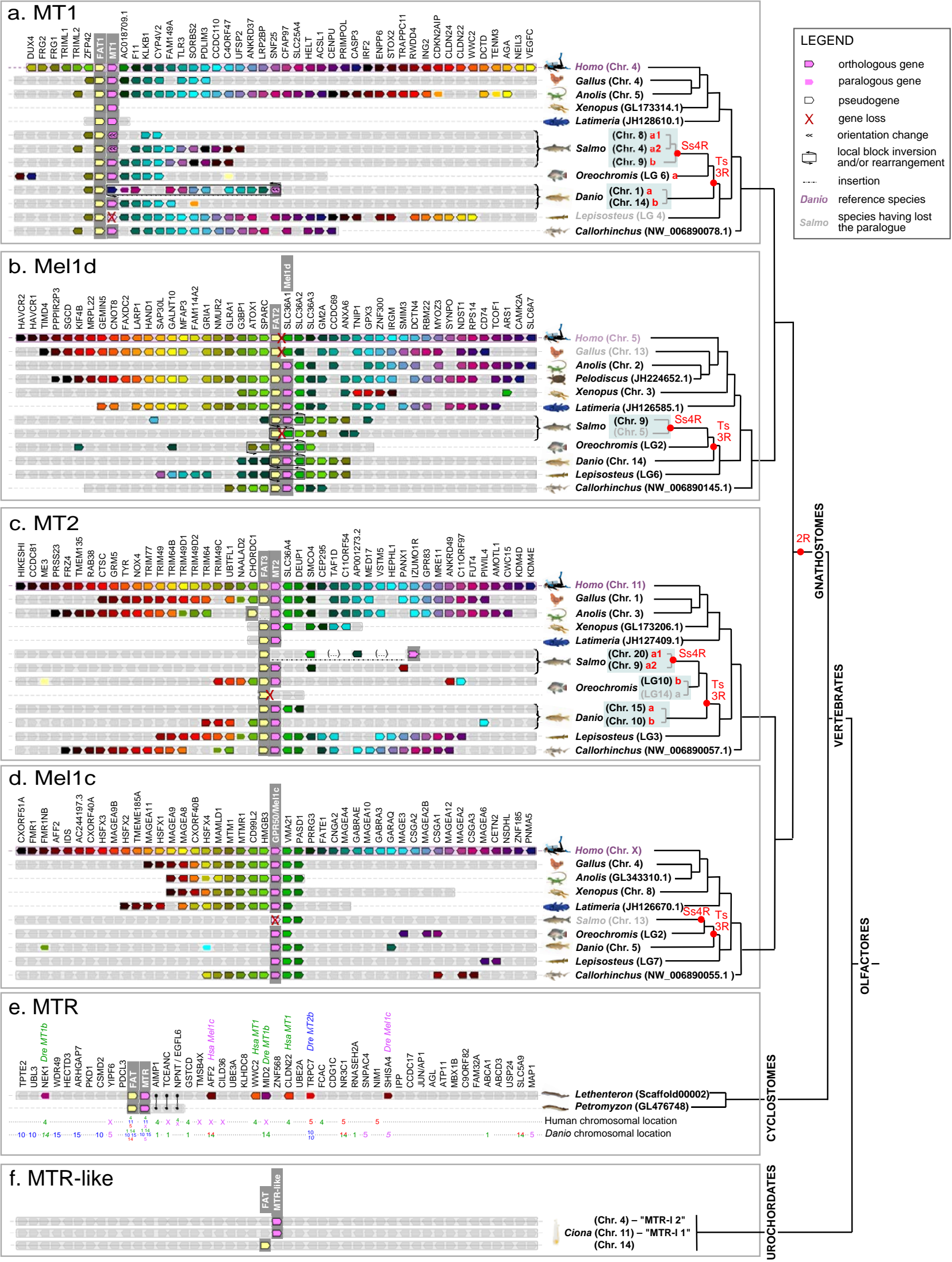


Mel1d

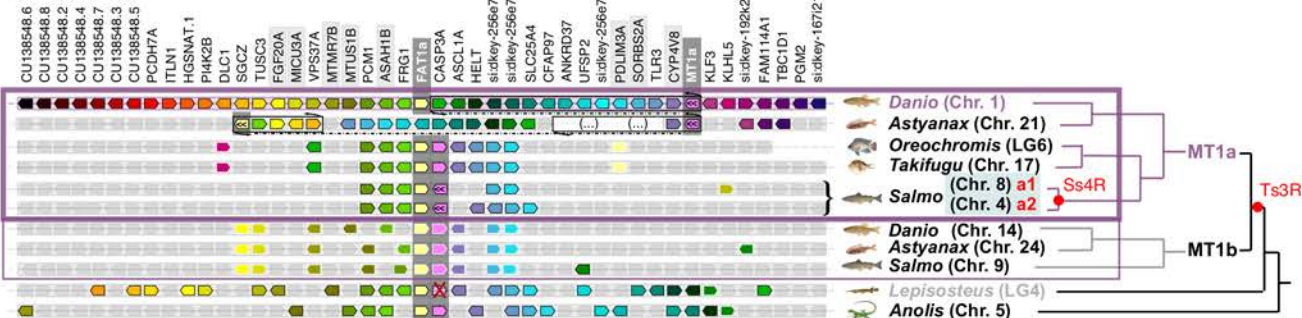
MT1/MTNR1A

MT2/MTNR1B

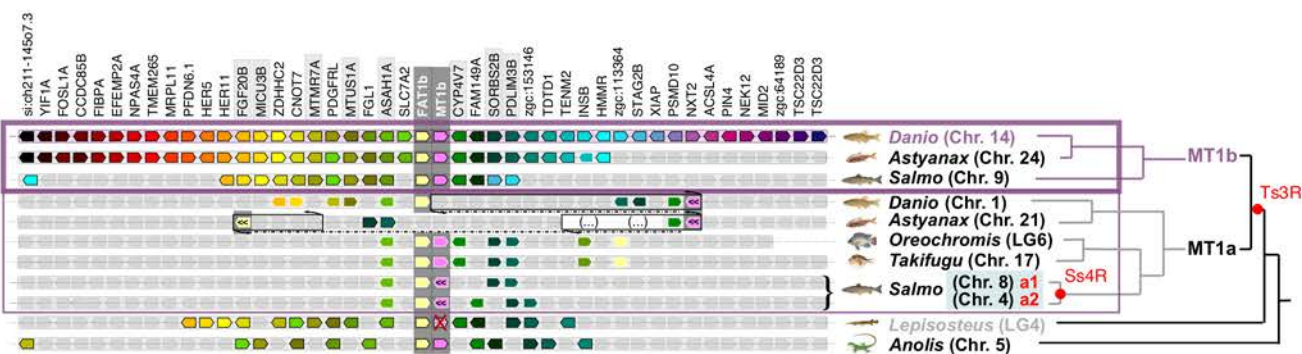
Mel1c



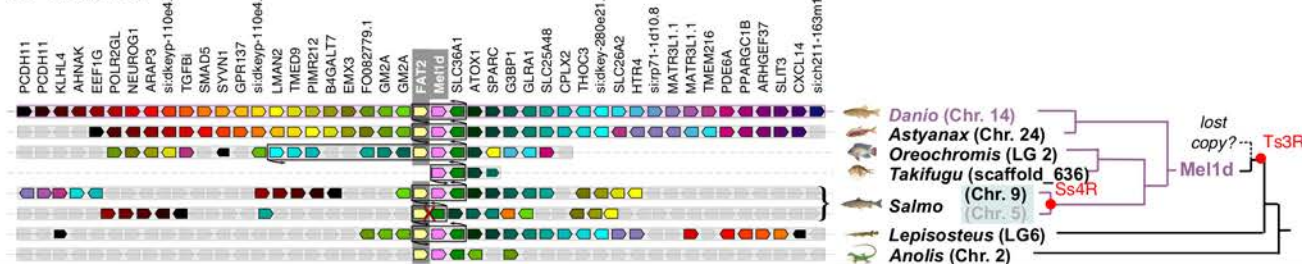
a. MT1a



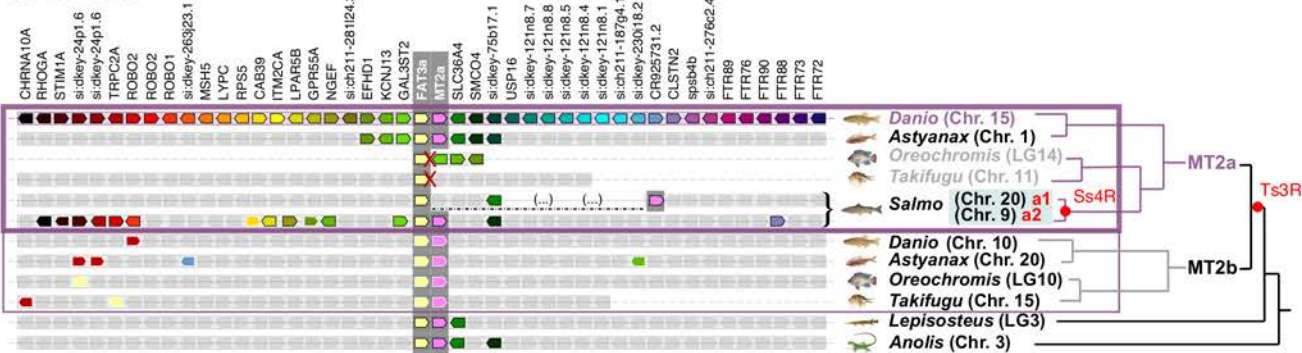
b. MT1b



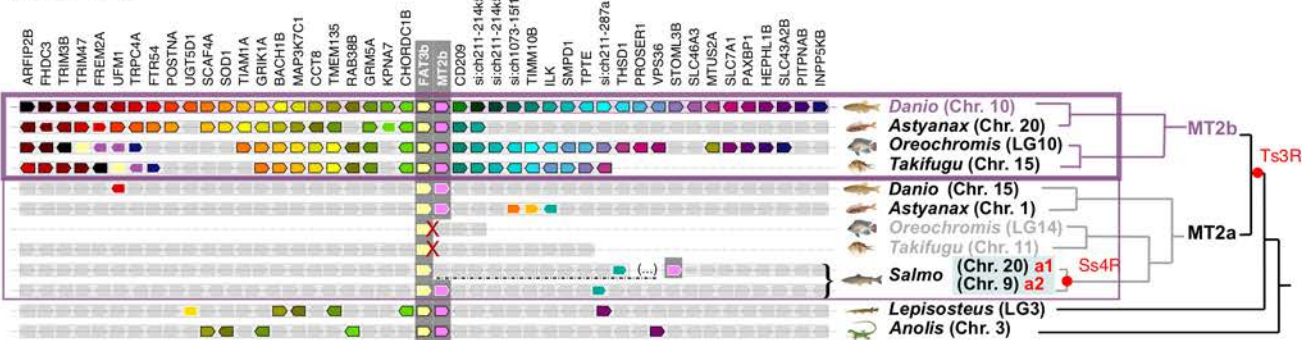
c. Mel1d



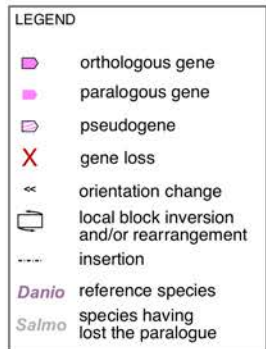
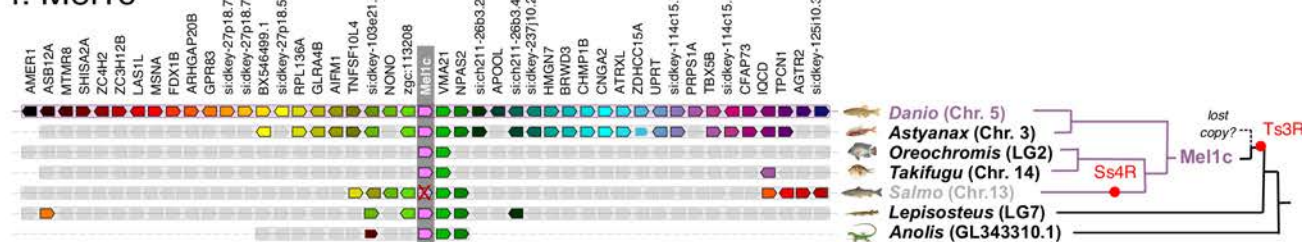
d. MT2a

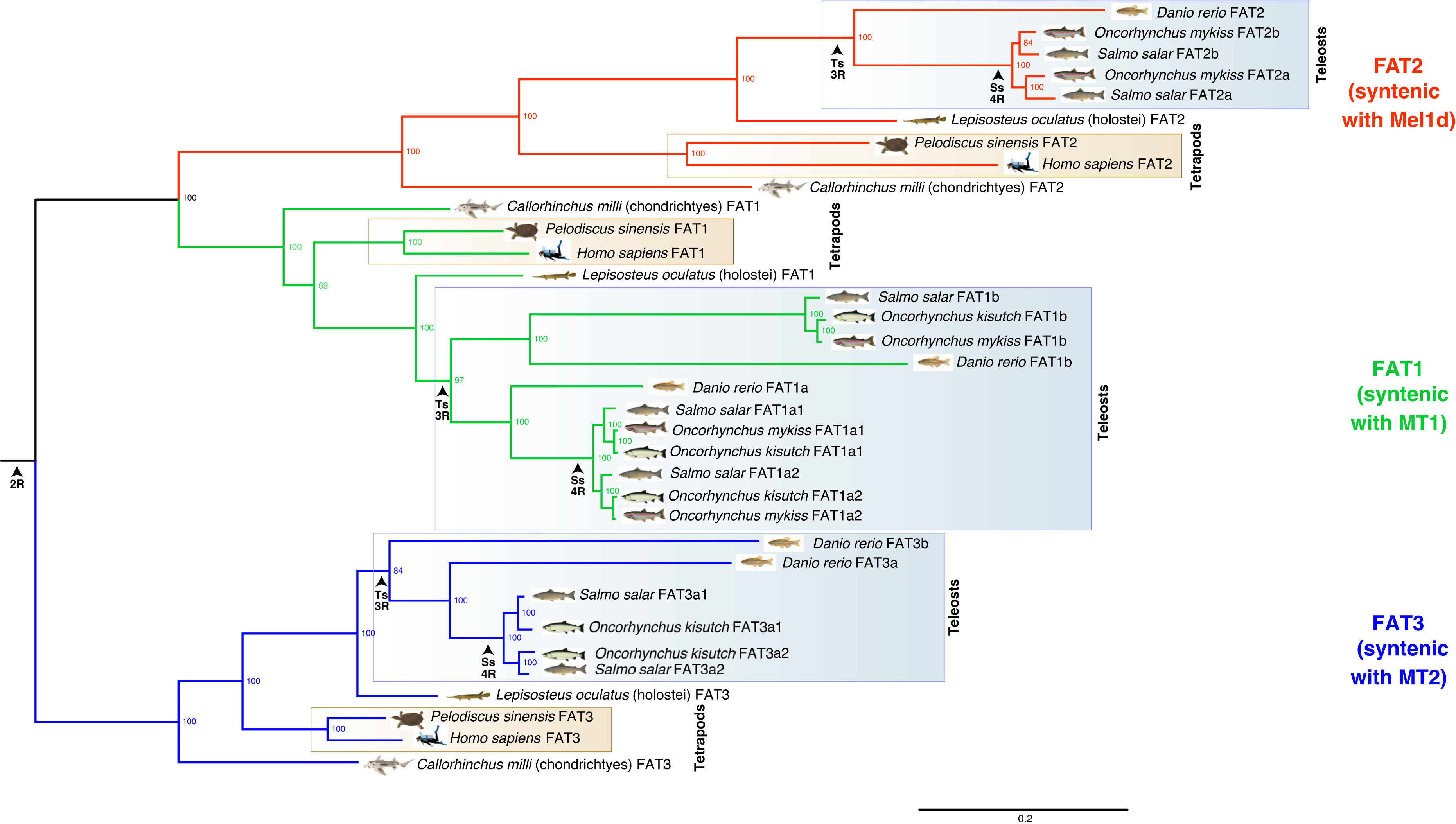


e. MT2b



f. Mel1c





Mutations affecting the function of hMe1d

Mutations affecting the function of hMe1T



Table 1. Phylogenetic assignment of teleost MTRs to a standardized nomenclature system.

Species		Receptors attributed from literature vs orthology group assignment from this study						References
		MT1		MT2		Mel1c	Mel1d	
		MT1a	MT1b	MT2a	MT2b			
Dr	<i>Danio rerio</i> (zebrafish)	Z1.7 (U31822.1)			Z2.6 (U31824.1)		Z1.4 (U31823.1)	Reppert <i>et al.</i> 1995(b)
		zMel1a1, Z1.7-4, mtnr1aa (NM_131393.1)	zMel1a3 (XM_688989.6)	zMel1b2, Z6.2, Mel1b-19, mtnr1ba (NM_131395.1)	zMel1b1, Z2.6-4, mtnr1bb (NM_131394.1)	zMel1c, Z2.3, mtnr1c (NM_001161484.1)	zMel1a2, Z1.4, mtnr1al (NM_001159909.1)	Shang & Zhdanova 2007
Om	<i>Oncorhynchus mykiss</i> (rainbow trout)	R1.7 (AF156262.1) = MT1a2					R1.4 (AF178538.1)	Mazurais <i>et al.</i> 1999
EI	<i>Esox lucius</i> (northern pike)			P2.6 (AF188871.1)			P1.4 (XM_010903666.1)	Gaildrat and Falcón, 2000
Oke	<i>Oncorhynchus keta</i> (chum salmon)	mel1a (AY356364.1) = MT1a2					mel1b (AY356365.1)	Shi <i>et al.</i> 2004
Sg	<i>Siganus guttatus</i> (golden rabbitfish)	Mel1a (DQ768087.1)				Mel1c (DQ768088.1)	Mel1b (DQ522314.1)	Park <i>et al.</i> , 2006, 2007a,b, 2014
Ca	<i>Carassius auratus</i> (goldfish)	Mel1a1.7 (AB378058.1)			Mel1b (AB378059.1)	Mel1c (AB378060.1)	Mel1a1.4 (AB378057.1)	Ikegami <i>et al.</i> 2009
		G1.7 (AB481372.1)		G6.2 (AB481374.1)	G2.6 type1 (AB481373.1)	Mel1c (AB481375.1)	G1.4 (AB481371.1)	Saito, unpublished
DI	<i>Dicentrarchus labrax</i> (sea bass)	dIMT1 (EU378918.1)			dIMT2 (EU378919.1)	dIMel1c (EU378920.1)		Sauzet <i>et al.</i> , 2008 Herrera-Pérez P <i>et al.</i> , 2010
Sse	<i>Solea senegalensis</i> (Senegal sole)				ssMT2 (FM213464.1)	ssMel1c (FM213465.1)	ssMT1 (FM213463.1)	Confente <i>et al.</i> 2010
On	<i>Oreochromis niloticus</i> (Nile tilapia)	mel1a (AY569971.1)						Jin <i>et al.</i> , 2013
Ec	<i>Epinephelus coioides</i> (orange-spotted grouper)	MT1 (JX524508.1)			MT2 (JX524509.1)			Chai <i>et al.</i> , 2013
Bp	<i>Boleophthalmus pectinirostris</i> (mudskipper)	Mtnr1a1.7 (KC622030.1)			Mtnr1b (KC622031.1)	Mtnr1c (KC622032.1)	Mtnr1a1.4 (KC622029.1)	Hong <i>et al.</i> 2014
Tn	<i>Takifugu niphobles</i> (grass puffer)	mel1a1.7 (AB492764.1)			mel1b (AB492765.1)	mel1c (AB492766.1)	mel1a1.4 (AB492763.1)	Ikegami <i>et al.</i> 2015
Pn	<i>Porichthys notatus</i> (plainfin midshipman - “singing” fish)	mtnr1A1.7 (HQ007044)		Mel1b (KT878765.1)			mtnr1a1.4 (HQ007045)	Feng & Bass, 2016, Feng unpublished
Amel	<i>Amphiprion melanopus</i> (cinnamon clownfish)	MT-R1 (HM107821.1)						Choi <i>et al.</i> , 2016

Phylogenetic assignment according to findings of this study; previous publications using distinct nomenclature systems are provided. Sequences in red signal a significant change in assignment.